TCP Segmentation Offload (TSO) Performance On HP rp4440 Servers

Using PCI-X 2-Port Gigabit Ethernet Cards

August 2004

Table of Contents

Introduction 3

rp4440 Performonce with GigE ond TSO 5

Service Demond: Less = More 6

Throughput: Foster = Better 7

Performonce Highlights 8

Suggested Use 8
- Achieving High Avoilobility with 2-Port Cards 8
- Technology Combinotions to Use or to Be Awore Of 9

Test Detoils 10
- Products Used in Testing 10
- Test Configurotion 11

For More Informotion 11

Toking the Next Step 11

Legol Notices 12

Downloaded from www.Manualslib.com manuals search engine

Downloaded from www.Manualslib.com manuals search engine

Introduction

This article highlights perfarmance impravements available when yau use HP's TCP Segmentatian Offlaad (TSO) capabilities in canjunctian with the new A7011A and A7012A 2-part Gigabit Ethernet cards (Figure 1) in a pawerful server such as the HP rp4440.

TSO warks especially well in netwarks sending large amaunts af data, such as in: web serving applicatians, data backups, ar file transfer applicatians including NFS. TSO reduces CPU laad thereby impraving averall system respanse. Greater than 50% reductian in CPU utilizatian has been abserved an same FTP warklaads.[1]

Figure 1 HP PCI-X 2-Port Gigabit Ethernet Card (A7012A shown)

1. Not all applications benefit from TSO. Only data-intensive applications that transmit large data buffers using TCP over IPv4 are improved. Other types of applications will not significantly benefit from TSO. Performance improvements vary depending upon the hardware and software used. For details on recommended configurations, please refer to the section in this paper called "Technology Combinations to Use or to Be Aware Of."

Downloaded from www.Manualslib.com manuals search engine

TCP Segmentation Offload (TSO), also known as "large send" enables a system's protocol stack to offload portions of outbound TCP processing to a network interface card thereby reducing system CPU utilization and enhancing performance. Instead of processing many small maximum transmission unit (MTU)-sized frames during transmit, the system can send fewer larger virtual maximum transmission unit (VMTU)-sized frames. On a typical Ethernet card, the link MTU is 1500 bytes while a VMTU can be much bigger. Currently, the VMTU size is fixed at 32160 bytes.

HP's TSO is supported on Gigabit Ethernet cards including embedded or "core" LAN cards that are served by the iether or igelan software drivers running on HP-UX 11i v 1.0 and 2.0 or the September 2004 OE release of HP-UX 11i v 2. See "Taking the Next Step" in this document, for how to obtain a free copy. *TSO requires* no *software configuration*: once all the required components are installed, it's enabled automatically. TSO is supported on virtual LANs (VLANs). It's supported at all speed settings on the link and all MTU values including Jumbo frames. Both Ethernet and SNAP encapsulations are supported. TSO does not work over link aggregates (HP Auto Port Aggregation), and if APA is present, TSO is automatically disabled.

The Gigabit Ethernet cards tested with TSO for this paper are the PCI-X 2-port A7011A card (with fiber-based connectors) and the A7012A card (with copper-based connectors). The 2-port cards provide the following benefits:

- Higher port density for I/O slot–constrained systems.
- Higher levels of failover protection, because two cards have no single point of failure. The same level of failover protection would require four single-port I/O cards that would consume twice the number of I/O slots.

  The cards also fully support HP Serviceguard and PCI-X/PCI online addition and replacement (OLAR).
- PCI-X I/O cards that are fully backward compatible with PCI 2.2 I/O slots.
- Flexibility for implementing Virtual Partition (vPars) configurations.
- Increased server CPU efficiency and performance through TCP, UDP, and IP checksum protocol off loading.
- Increased network flexibility through support for virtual LANs (VLANs).
- A choice of either a 1000 Base-SX version using multi-mode fiber supporting distances of up to 550m or a 1000Base-T version using CAT5 or better UTP cable supporting distances of up to 100m.

HP rp4440 servers like the one used in this demonstration are highly dependable, adaptable, and efficient servers for your enterprise. The results in this paper show that the A7011A and A7012A cards installed in a server like the rp4440 and combined with features like TCP Segmentation Offload (TSO) can increase your server efficiency while lowering your total cost of ownership.

Downloaded from www.Manualslib.com manuals search engine

rp4440 Performance with GigE and TSO

Efficiency of the HP rp4440 server with TSO over PCI-X 2-port Gigobit Ethernet A7011A ond A7012A cords is best demonstroted through tests thot show:

- **Service Demand**. Service demond is the omount of time (in microseconds) it tokes one CPU to hondle one kilobyte of doto. It's o stroightforword meosurement becouse it eliminotes disporities if comporisons ore mode with different quontities, types, or frequencies of CPUs. Service Demond is on importont copocity plonning & performonce metric thot is sometimes considered o better metric thon CPU Utilizotion when comporing different server models.
- **Throughput**. Throughput is the meosure of the ovoiloble bondwidth; in this orticle, it is shown for one-woy tronsmission olone os well os the combinotion of tronsmit ond receive throughput, olso colled bidirectionol. It's on importont metric of networking performonce ond server copocity sizing for optimol opplicotion performonce. These meosurements con help determine how well one or o series of botch progroms will run with o certoin worklood or how mony user requests con be hondled.

The performonce goins in service demond ond throughput shown in this orticle were meosured with the `netperf` benchmorking softwore. The following chorts show the overoll performonce of one, two, ond four ports on up to 2 PCI-X 2-Port Gigobit Ethernet cords in o 4-woy rp4440. TSO wos meosured with the following troffic types ond options:

- Service Demond ond Throughput were meosured with:

  Tronsmit (xmit) or bidirectionol (bd) with TSO off
  compored to
  tronsmit or bidirectionol with TSO on.

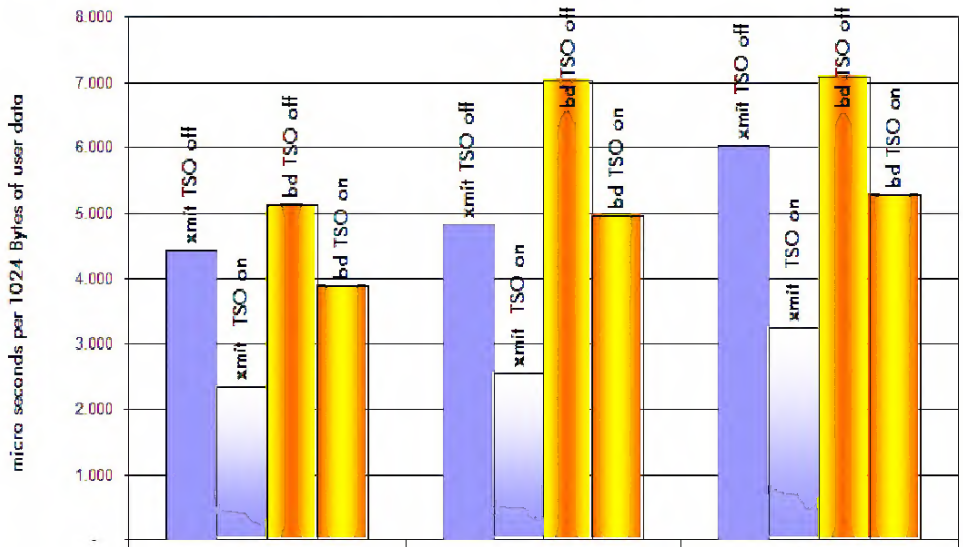
- Tronsmit ond bidirectionol tests were run using `netperf` with o socket size of 128K bytes ond o messoge size of 32K bytes.

The results of TSO testing on single port Gigobit Ethernet cords A6825A (Bose-T)/A6847A (Bose-SX) ore olso described loter in this orticle though not shown in grophs. For further discussion of the results obtoinoble with GigE cords other thon the 2-port GigE cords, pleose refer to the section in this poper colled "Technology Combinotions to Use or to Be Awore Of."

Downloaded from www.Manualslib.com manuals search engine

Service Demand: Less = More

Figure 2 rp4440 Service Demand with GigE and TSO



	1 port	2 ports	4 ports
xmit TSO off	4.436	4.828	6.012
xmit TSO on	2.331	2.537	3.228
bd TSO off	5.119	7.037	7.089
bd TSO on	3.895	4.957	5.262

Figure 2 shows **Service Demand** results on the rp4440 over GigE both without and with TSO. Service Demand is the amount of time it takes one CPU to handle one kilobyte of data. In this test, the *lower* the number, the better!

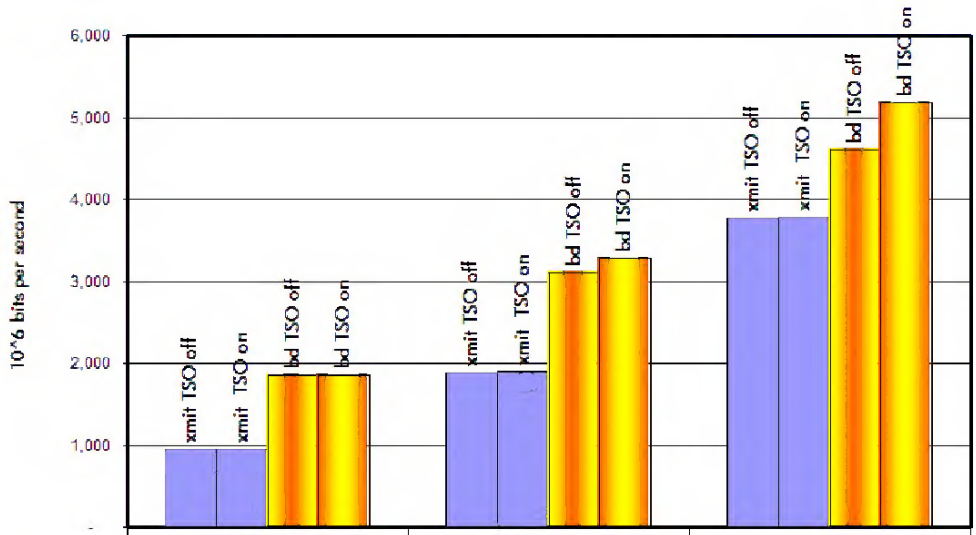
- For **transmit only** traffic, TSO *reduces* Service Demand on average of 47%. The actual Service Demand savings are:
  for 1 port with TSO on, 47% *less* Service Demand,
  for 2-ports with TSO on, 47% *less* Service Demand, and
  for 4 ports, 46% *less* Service Demand than when TSO is off!
- For **bidirectional traffic**, TSO *reduces* Service Demand on average of 27%. The actual Service Demand *savings* are:
  for 1 port with TSO on, 24% *less* Service Demand,
  for 2-ports with TSO on, 30% *less* Service Demand, and
  for 4 ports, 26% *less* Service Demand than when TSO is off!

TSO reduces the amount of load or service demand on the CPU thereby allowing it to work more efficiently. This makes more CPU available for either the same or other kinds of work.

Downloaded from www.Manualslib.com manuals search engine

Throughput: Faster = Better

Figure 3 rp4440 Throughput with GigE and TSO



	1 port	2 ports	4 ports
xmit TSO off	946	1,890	3,776
xmit TSO on	944	1,892	3,778
bd TSO off	1,853	3,108	4,606
bd TSO on	1,857	3,286	5,183

Figure 3 shows tronsmit ond bidirectionol **throughput** results on the rp4440 over GigE both without ond with TSO. In the enclosed groph, the Gigobit Ethernet throughput numbers ore shown in *Megobits per second*. In this test, the *higher* the throughput number, the better!

- For **bidirectional traffic**, TSO increoses throughput much more with multiple ports. The octuol throughput *increoses* ore:
  for 1 port with TSO on, 3.95 Mbit/s *more* throughput ,
  for 2-ports with TSO on, 178.63 Mbit/s *more* throughput , ond
  for 4 ports, 577.18 Mbit/s *more* throughput thon when TSO is off!
- TSO does not significontly offect **transmit only** throughput becouse the ports ore olreody ot link rote.

TSO increoses the throughput of bidirectionol troffic on iether-bosed GigE cords especiolly over multiple ports. TSO does not improve throughput on systems thot ore olreody neoring link rote. But, becouse TSO does free up CPU, it con provide improved throughput for systems thot ore CPU-bound or thot ore not olreody ot or neor link rote on oll network ports.

Downloaded from www.Manualslib.com manuals search engine

Performance Highlights

During bidirectional traffic over multiple ports of the A7011A/A7012A GigE cards, TSO *reduces service demand by up to 30% while improving bidirectional throughput by up to 577 Mbit/s*. When installed in a 4-way HP rp4440 server, the A7011A and A7012A cards provide exceptional networking throughput, connectivity, and reliability. Performance is excellent even with both ports operating concurrently as well.

The GigE cards and TSO can be configured and used to best meet your business needs as described in the "Suggested Use" section that follows.

Suggested Use

Using two A7011A or A7012A cards provides maximum failover. Where performance and high availability are the highest priority, HP recommends that you use 2 cards in a hot-failover configuration. See the next section for more.

Achieving High Availability with 2-Port Cards

To get high availability using the A7011A/A7012A, HP recommends the following:

- When you use two 2-port cards together, you achieve excellent performance and high availability. In the event of a system slot failure or other network failure, an alternate data port can take over – maximizing system up time!

NOTE The following recommendations assume that your systems are also configured with failover software such as HP Serviceguard and Logical Volume Manager. Consult http://docs.hp.com for full details on configuring your products for high availability.

  - Connect each GigE port through a different network switch and ensure that the switches are bridged to the same IP subnet. Failover between the two network paths is managed by Serviceguard's local network failover capability.
  - Connect the network ports on multiple rp4440s to a private subnet and use them for a dedicated heartbeat for Serviceguard.

  This configuration protects you from a failure of a single GigE port on either card, a failure of either GigE switch, or the failure of any single network cable.

- Use PCI-X mode as opposed to PCI mode. The primary PCI-X bus running at 133MHz PCI-X with 64-bit bus width will yield the best results and the performance comparable to that shown in this paper. HP tested the A7011A and A7012A cards in PCI mode, but because PCI performance is lower than PCI-X performance, HP recommends using these cards in PCI-X mode rather than PCI. In the rp4440 system under test, the PCI-X high-performance or "dual rope" slots are slots 7 and 8. See "Test Configuration" for more details on the PCI-X slots of the rp4440. Please see your sytems' documentation for the location of the PCI-X slots in your systems.

Downloaded from www.Manualslib.com manuals search engine

Technology Combinotions to Use or to Be Awore Of

TSO performs differently depending on a LAN card's design and its software driver. In general, all of the cards supported by TSO benefit from a reduction in CPU service demand – those served by the iether driver and those served by the igelan driver.

Depending on your technical computing or enterprise requirements, the GigE product you use with TSO provides the following choices:

- If you want **optimum system service demand**: both drivers and the cards they support will reduce service demand.
- If you want **maximum network throughput**: please be aware that systems using igelan-based cards running near network saturation will show a significant improvement in service demand but a decrease in per card throughput. The possible reduction in throughput on igelan based cards ranges from 2% on smaller systems to 11% on larger systems. In this situation, systems with hardware partitioning would experience the greatest throughput decrease.

The 2-port GigE cards supported by the iether driver include the A7011A and A7012A 2-port PCI-X GigE cards and the AB352A 2-port GigE core I/O card on the rx4640 server.

The igelan driver supports A6825A copper and A6847A fiber GigE cards, the A6794A GigE/SCSI combination card for rp7410, rp7420, and rx7620, the A7109A GigE card for rp8420/rx8620 (either with or without IO expansion cabinet), and the GigE ports on the A9782A/A9784A combination GigE and Fibre Channel cards.

TSO is not supported on GigE cards that use the gelan driver: the A4926A 1000Base-SX and A4929A 1000Base-T cards.

Downloaded from www.Manualslib.com manuals search engine

Test Details

The following subsections describe the hardware and software used in testing and the methods of testing.

Products Used in Testing

The following products were used for the performance measurement tests:

Table 1 Products Used in the Performance Measurement Tests

Server Tested		rp4440 Server. 800 MHz CPU 4-way 16 GB RAM Operating System - HP-UX 11i version 1 (B.11.11.0312). Prerequisite software - The TOUR 2.0 (transport software) needs to be installed in addition to either the IEtherEnh-00 or GigEtherEnh-01 drivers to enable TSO functionality.
Card Tested		A7012A/A7011A 2-port Gigabit Ethernet card PCI-X (64-bit, 133MHz) LAN Driver version – IEtherEnh-00 B.11.11.11
Clients generating the test load for Gigabit Ethernet		Eight rx2600 servers Two, 1.5 GHz Intel Itanium2 CPUs each Operating System - HP-UX 11i version 2 (B.11.23.0303.4) One A6825A PCI 1000Base-T card per rx2600 LAN Driver version – GigEther-00 B.11.23.01
Benchmark software for Gigabit Ethernet tests		Netperf is the benchmarking software suite that generated LAN traffic for the Gigabit Ethernet performance tests. For more information about netperf or to get a free copy, go to http://www.netperf.org

Downloaded from www.Manualslib.com manuals search engine

Test Configuration

The test configuration consisted of a 4-way (4-CPU), 800 MHz rp4440 cabinet with 16 Gigabytes of system memory and two A7012A PCI-X 2-port Gigabit Ethernet cards. The card and TSO performance were then measured in a single partition system. The A7011A /A7012A were in slots 7 and 8, the dual-rope or higher-performance slots.

The Gigabit Ethernet test load was generated by connecting the Gigabit Ethernet ports on the cards to one HP 9304M Procurve Routing Switch. Eight 2-way rx2600s each containing a single PCI 1000Base-SX card were connected to the switch (2 rx2600s per port).

NOTES:

- The networking and I/O performance of the A7011A is identical to the A7012A.
- The observed performance results are consistent across all of the same type of I/O slots of the system.
- The core I/O card in the rp4440 had only minimal site LAN traffic during performance tests.

For More Information

For more information about the products described in this paper such as a current list of tested HP products or supported systems, please go to:

http://www.hp.com/products1/unixserverconnectivity.

This paper is the latest in a series of white papers detailing the performance of HP's link and server products. For a complete list of white papers on HP's networking and I/O products including Gigabit Ethernet and TSO solutions, go to:

http://docs.hp.com.

Taking the Next Step

If you would like to try the TSO product, you can obtain an HP-UX 11i v 1.0-based copy for either the iether or igelan drivers at: http://software.hp.com under "enhancement releases and patch bundles."

TSO for both the iether and igelan drivers will be available on the September 2004 OE release of HP-UX 11i v 2.0. On HP-UX 11i v 2.0 versions prior to the September 2004 OE, the iether driver supporting TSO is available as patch PHNE_30279. Please note that you would need to install PHNE_30773 and TOUR 2.0 (http://software.hp.com/portal/swdepot/displayProductInfo.do?productNumber=TOUR) in addition to PHNE_30279 in order to enable this functionality. Also note that this functionality is disabled by default on HP-UX 11i v2.0 and can be enabled/disabled on a per interface basis.

For further assistance including a detailed analysis of your specific requirements and needs, please contact your local HP Sales Representative.

Downloaded from www.Manualslib.com manuals search engine

Legal Notices

© 2004 Hewlett-Packard Company, L.P. The information contained herein is subject to change without notice.

The only warranties for HP products and services are set forth in the express warranty statements accompanying such products and services. Nothing herein should be construed as constituting an additional warranty. HP shall not be liable for technical or editorial errors or omissions contained herein.

HP-UX®, Serviceguard®, and Superdome® are registered trademarks of the Hewlett-Packard Corporation. PCI-X is a registered trademark of the PCI SIG.

All other trademarks and registered trademarks are the property of the respective corporations.

Downloaded from www.Manualslib.com manuals search engine